NetCDF-4
The Marriage of Two Data Formats

Ed Hartnett,

Unidata

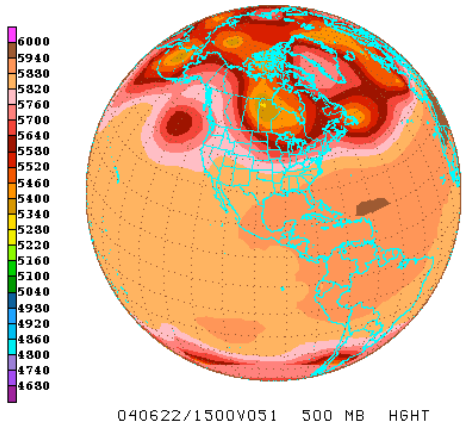
June, 2004

NetCDF-4 Project

- Seeks to merge netCDF and HDF5 data formats.
- Provides netCDF features for HDF5 users.
- Provides HDF5 features for netCDF users.
- On schedule for release in March, 2005.
- Joint project of Unidata and NCSA.

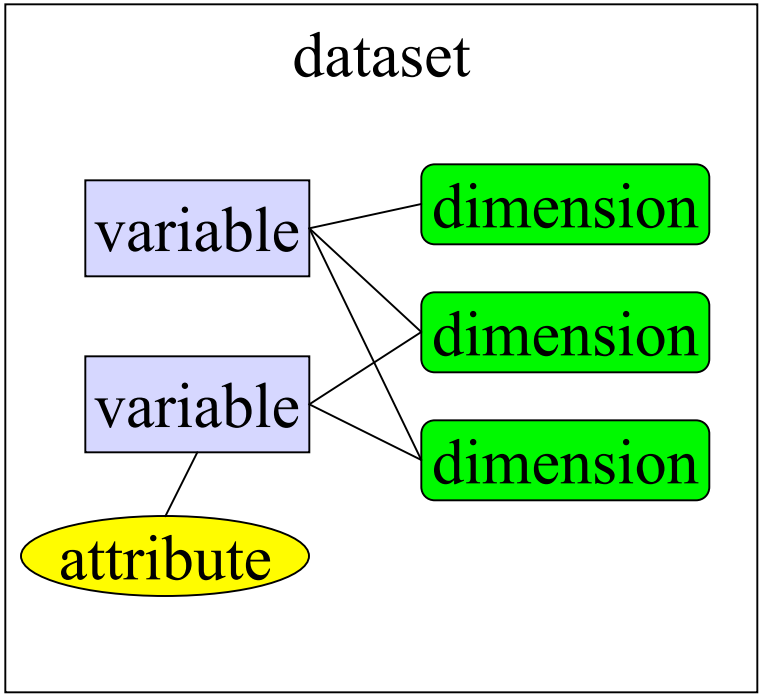
What is Unidata?

- Unidata is a diverse community of institutions vested in the common goal of sharing data, and tools to access and visualize that data.



What is NetCDF?

- A straightforward scientific data model.
- Programming APIs in C, Java, Fortran 77/90, C++, Perl.

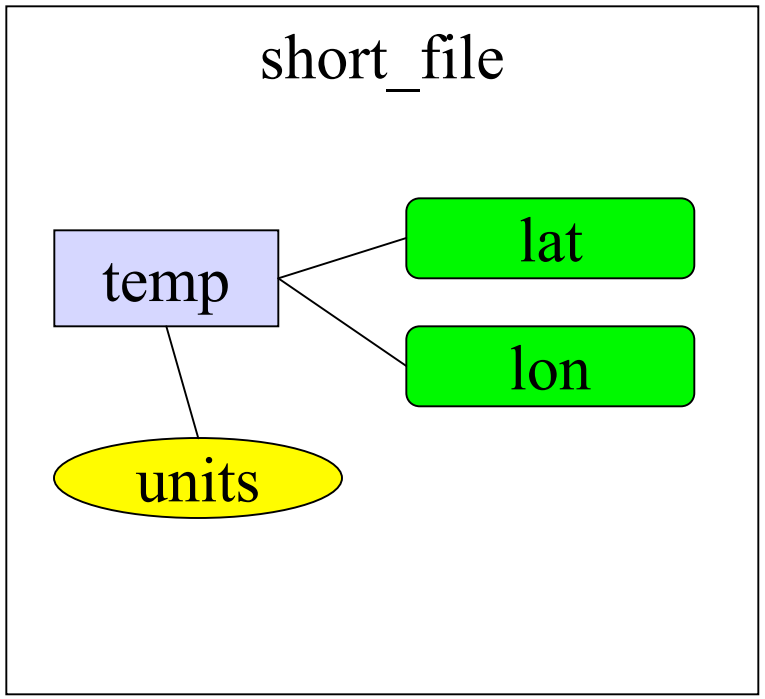


NetCDF C Example

```
/* Create new netCDF file. */
   nc_create("short_file.nc", NC_CLOBBER, &ncid);

   /* Define metadata. */
   nc_def_dim(ncid, "lat", LAT_LEN, &dimids[0]);
   nc_def_dim(ncid, "lon", LON_LEN, &dimids[1]);
   nc_def_var(ncid, "temp", NC_FLOAT, NDIMS, dimids, &varid);
   nc_put_att_text(ncid, varid, "units", strlen(CELSIUS), CELSIUS);
   nc_enddef(ncid);

   /* Write data. */
   nc_put_var_float(ncid, varid, (float *)data);

   /* We're done! */
   nc_close(ncid);
```

Example in CDL Notation

```
netcdf short_file {
dimensions:
      lat = 3 ;
      lon = 2 ;
variables:
      float temp(lat, lon) ;
            temp:units = "celsius" ;
data:

 temp =
  10, 10.1,
  10.2, 10.3,
  10.4, 10.5 ;
}
```



Why is NetCDF Popular?

- Longevity - First developed by Glenn Davis of Unidata in 1988.
- Ubiquity - NetCDF has been ported to many platforms, tools, and programming languages.
- Simplicity – Scientists/programmers can work with netCDF immediately.

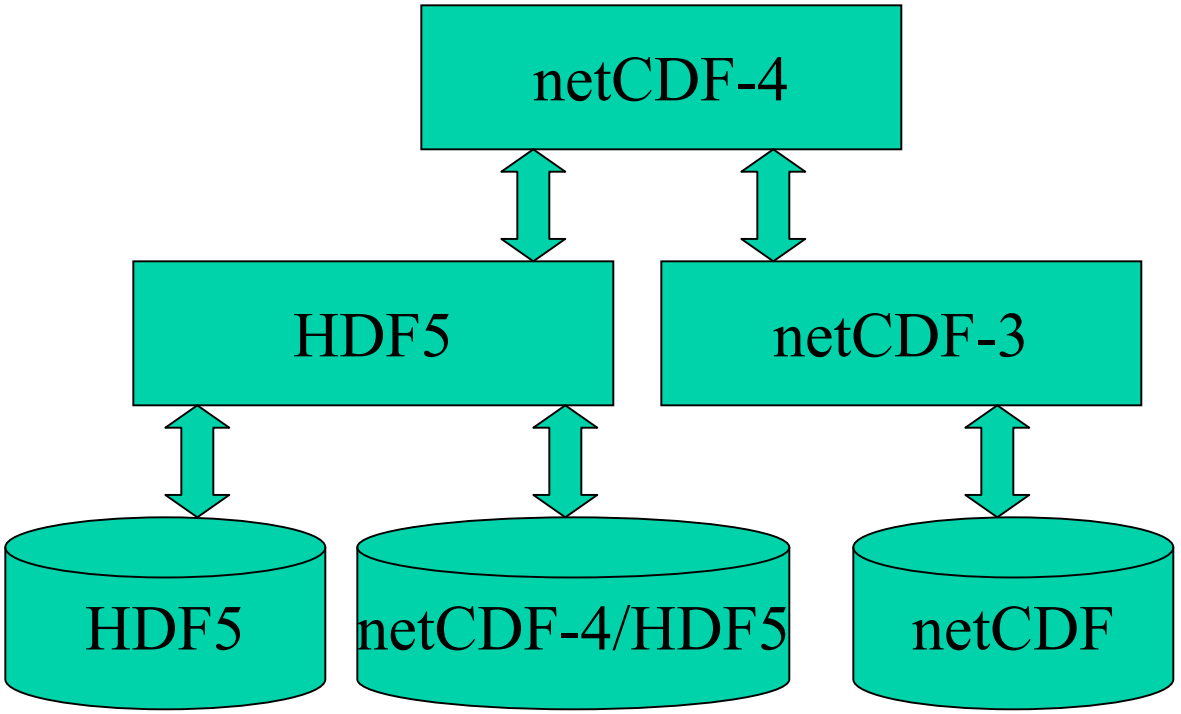
NetCDF Limitations

- Size – files larger than 2 GB are tricky.
- Dimensionality – only one unlimited dimension is permitted in netCDF.
- Interoperability – netCDF data files are completely incompatible with HDF files.
- Organization – data model not well-suited for hundreds or thousands of variables.

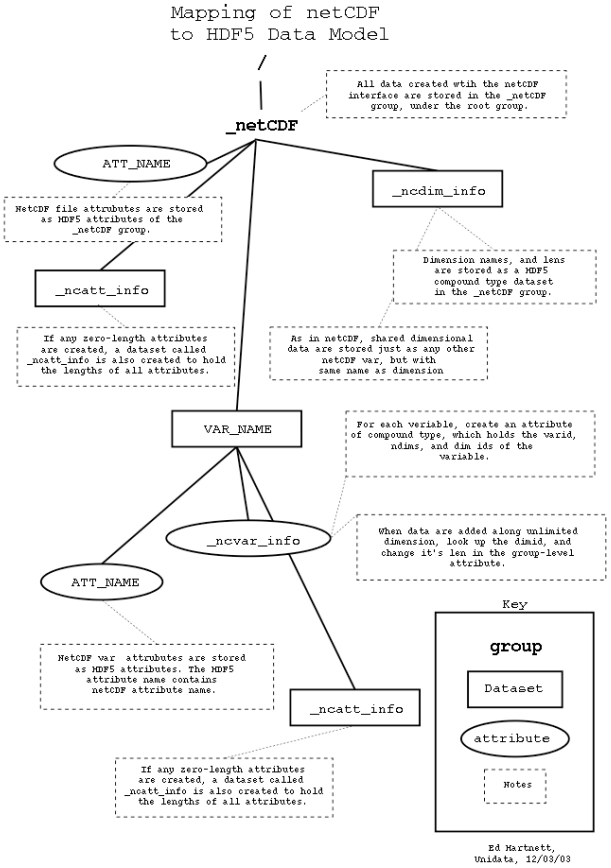
NetCDF-4

- Uses HDF5 as back-end storage.
- Extends the API to support some new features.
- No size and dimensionality restrictions.
- Additional organization and access to advanced HDF5 features.
- HDF5 users see netCDF-4 as a new interface.

NetCDF-4 Architecture



NetCDF-4 Dataset Organization



- NetCDF-4 data exist under a group called "_netCDF".
- HDF5 datasets and attributes are used to store netCDF metadata.
- NetCDF variables are stored as HDF5 datasets.

NetCDF-4 Prototype

- Passes netCDF test suite.
- Delivers acceptable performance.
- Complete implementation of netCDF-3 API.
- Demonstrates backward API and file format compatibility.

Schedule

- Alpha release of netCDF-4 in October, 2004.
- Beta release of netCDF-4 with netCDF v. 3.7.1, January, 2005.
- NetCDF-4 release with netCDF v. 4.0 in March, 2005.

Web Resources

- Unidata - www.unidata.com
- NetCDF - www.unidata.ucar.edu/packages/netcdf
- NetCDF-4 - www.unidata.ucar.edu/packages/netcdf/netcdf-4

Contacts

- PIs – Russ Rew and Mike Folk
  russ@unidata.ucar.edu,
  mfolk@ncsa.uiuc.edu
- Unidata Programmer – Ed Hartnett
  ed@unidata.ucar.edu

Conclusion

- NetCDF-4 is funded by NASA's Earth Science Technology Office.
- Unidata is sponsored by the NSF.
- NCSA is funded by NASA, NSF, and DOE.